Response to: anonymous reviewer(s) two (cpd-11-C2884-2016)

General comments:

GC1: *The paper presented by Grimm and Potts is a complete evaluation of the limitations associated with the Coexistence Approach (CA). While most of the critics are well argued, the tone of the paper is more confrontational and aggressive than the subject would actually deserve, which becomes irritating for the reader before the end of the introduction.*

R1: Please see response **R3** in **Response to Reviewer 1** regarding the confrontational nature of the title (and general tone).

GC2: *Most of the theoretical aspects discussed in this paper are valid; however, considering the conclusions of the paper "we suggest that the method be discontinued and that all past reconstructions be disregarded and revisited using less fallacious methods", (potential) alternatives were expected by the end of the discussion.*

R2: The aim of this paper is to stop the use of an obviously flawed method. Other methods exist, some of which are cited in the text, but a method comparison is not relevant here. If a method produces a result that is in all likelihood highly erroneous, then it simply should not be used. However, we have included a new section (section **5**) entitled "**Where to go from here?**" which does provide potential alternatives and suggestions for the community to explore.

GC3: *Repeatedly, the authors point to aspects of the method that they consider to be critical (cf. below) but they never propose any guidance to overcome them. Without those, the paper doesn't bring anything new that hasn't been previously discussed in the original publications (Mosbrugger and Utescher, 1997; Utescher et al., 2014).*

R3: Please see **R2**. Some of these points were indeed mentioned in Mosbrugger and Utescher (1997), who do not provide any discussion, and are only superficially touched on in Utescher et al. (2014). The main fallacies (**Table R1** in **Response to Reviewer 1**) are not addressed or recognised. Advances in palaeoclimate reconstruction methods for Eurasian fossil assemblages has largely stalled due to the presence of the Coexistence Approach (CA) following a “don’t fix something that produces something” thinking (see **GC6**). Once researchers are able (allowed) to look beyond CA, we are certain that other, already available taxon-based methods (see text) will be used instead, and, ideally, validated regarding the principal question: is it possible to reconstruct precise palaeoclimates with sufficient accuracy using a nearest-living-relative, taxon-based approach?

> ***GC3, example A****: P5731 L2: "In addition, the degree of systematic relatedness of a fossil to an NLR requires the placement of fossils within a phylogenetic framework (i.e. a tree or network) and this has never been explored in any Coexistence Approach study."*
>
> **R3a:** There is no need to overcome this, since Assumption 1 of the Coexistence Approach is superfluous (see Fig. 2 demonstrating the difference between the purported theory of Coexistence Approach and the principle of physiological uniformitarianism behind the application of nearest-living-relative approaches in general).
>
> ***GC3, example B:*** *P5732 L15: "Novel procedures and methods are required that take cognisance of the fact that the NLR niche is likely to be far broader than can be expected for that of the fossil."*
>
> **R3b:** This *is* a proposition. It is beyond the scope or aims of this paper to propose novel methods.
>
> ***GC3, example C:*** P5742 L24: "[. . .] we suggest that all palaeoclimate reconstruction studies using the Coexistence Approach be disregarded and that the palaeoassemblages be revisited with improved methods and careful, well-documented, and well-investigated NLR-associations."
>
> **R3c:** This, again, *is* a proposition.

GC4: *The paper lacks of a clear discussion on how to improve the situation. In the realm of palaeoclimatic reconstructions, it is largely recognised that none of the existing climate reconstruction method is ideal and performs well in every situations.*

R4: Please see **R2**.

GC5: *The authors recommend that the community stops using the CA but without proposing any alternative.*

R5: Please see **R2**.

GC6: *The CA, with all its flaws and limitations, at least allows the production of data and, hence, initiate the beginning of a discussion. Do the authors suggest that palaeoclimatic reconstructions should be put on hold until further notice?*

R6: No. After all, there is a plethora of other methods, and none of them have ever been applied on the many Eurasian floras studied exclusively with the Coexistence Approach (CA). The *results* of CA are likely to be, at best, dubious, and, at worst, entirely wrong. How can they form a basis for discussion, when the famed precision is directly related to errors of practical (Grimm and Denk, 2012; Grimm et al., 2015) or theoretical (this paper: Grimm and Potts, 2015) nature? Which is better for investigating the past: highly precise but entirely wrong climate reconstructions (which would mislead any narrative) or a phase of rebuilding and reinvestigation of the available and novel methods? We argue that the latter is necessary. Also, please see **R2.**

GC7: *Additionally, many of the criticisms are in fine more related to the application of the method rather than to the method itself, making them therefore more practical/methodological than theoretical. The title should be adapted.*

R7: Please specify, which criticism applies only to application of the method but not the method itself and where to draw the line between practical and theoretical aspects of a method.

Specific comments:

SC1: *§1 Introduction:*

Only the mutual climate range method is considered in the introduction. Even if the method is still currently the most widely used and certainly the object of the paper, references to the recent efforts made to improve those methods – specifically with the development of the methods based on probability density functions (pdfs) (Chevalier et al., 2014; Gebhardt et al., 2007; Harbert and Nixon, 2015; Kühl et al., 2002; Truc et al., 2013) – should be explicitly mentioned here and not only at the end of section 2.2 on page 5737, especially since they do use more information than just the extrema and central tendency measures (one of the limitations of MCR highlighted by the authors).

R8: Please see **R2.** In addition, we are uncertain as to why these should be mentioned earlier.

SC2: *§2.1 Assumptions*

"The first assumption has never been used in the application of the Coexistence Approach, and the three others superimpose additional uncertainty on the method and are easily violated, particularly if the aim is high accuracy and precision." The end of this sentence seems largely dishonest. Nobody claims that high accuracy and precision are achievable goals with such methods and that for many reasons (e.g. the uncertainties inherent to the input variables (fossils and calibration data) as discussed in Utescher et al. (2014)). This is even truer for climate reconstructions of the Miocene.

R9: A measurement system is considered valid if it is *both* accurate *and* precise (Wikipedia - 2015 - Accuracy and Precision). It is absolutely true that (Utescher et al., 2014) avoid providing any value for the assumed (or hoped) accuracy of Coexistence Approach reconstructions. Regarding precision, they say that it lies at 2.1 °C at mean for mean annual temperature (see also **R3** and **Table R1** in **Response to Reviewer 1**)

That nobody claims high accuracy and precision is, however, in stark contrast with actual publications involving the same authors, particular with respect to the Miocene. By reporting highly precise results and discussing these *as is*, assumes that they are also (highly) accurate. There is also a question of semantics here: how can a method that is *not* accurate nonetheless be described as "*robust*" (Utescher et al., 2014) and "*reliable*" (http://www2.geo.uni-bonn.de/Palaeoflora/texte/coexistence_approach.html)?

Below are some examples from text extracted only from the 2015 website or publications using the Coexistence Approach featuring one or several authors of the 2014 publication (see Grimm and Denk, 2012; Grimm et al., 2015, for more examples):

> "*The reliability and resolution of the coexistence approach are tested with various techniques* [not further specified or documented] *and proved to be very good; for instance, the resolution* [i.e. precision] *with respect to the mean annual temperature can be up to 1° C.*"
>
> (http://www2.geo.uni-bonn.de/Palaeoflora/texte/coexistence_approach.html)

> "*Mean annual temperatures (MATs) were **around 1 °C higher** than today, suggesting a cooling trend since the late Pliocene.*" (Huang et al., 2015; paper coauthored by three authors of the 2014 paper including the first author). Such as statement would be impossible without believing, unconditionally, in *very high* accuracy *and* precision.

> "*In the CA, at least 10 NLR taxa contributing with climate data are required to obtain reliable results (Mosbrugger and Utescher, 1997)....As regards the MAT, the **mean precision** of the results, i.e. the mean width of the coexistence intervals **amounts to 1.7 °C (std. 1.3 °C), and to 143 mm (std. 75 mm)** for the MAP, respectively. Thus, data obtained for the megafloras of Primory'e are well within the usual uncertainty range of CA data (Utescher et al., 2014)* [only data for MAT given in the cited paper based on possibly erroneous reconstructions; see **R3** in **Response to Reviewer 1**] … *The MMCO* [mid-Miocene climate optimum] *is mainly evident from MAT and WMMT in the Primory'e record. The data obtained for the Langhian Novokachalinskaya 9151 flora indicate a **moderate warming by 1–2 °C** (MAT, WMMT).*" (Utescher et al., 2015)

SC3: *§2.3 Pseudo-logic:*

"Ignoring these logical inconsistencies in the conception and application of the method, the Coexistence Approach still cannot be expected to reproduce a robust quantitative

reconstruction of the palaeoclimate, as (1) assumptions are likely to be violated but cannot be detected, (2) one cannot avoid using higher-level taxa to represent fossil species or morphotypes, and (3) the fossil record will always be incomplete to different degrees, and this will affect the calculated coexistence interval." As any reconstruction method, the CA has its flaws. Many of the points raised here are, however, duly acknowledged in the original publication (Utescher et al., 2014).

R10: The flaws of the Coexistence Approach (CA) are **fundamental**, ***both*** in theory (Grimm and Potts, 2015) and practise (Grimm and Denk, 2012; Grimm et al., 2015). Point *(1)* is not really addressed following the wrong belief of the original 1997 publication that violations will surface via "climatic outliers" (see Section 3.1 *The concept and power of outliers* in Utescher et al., 2014). Point *(2)* is not even mentioned as an issue, although being rich of discussion in case of other approaches (e.g. Boyle et al., 2008 and literature cited therein). Regarding Point *(3)*, the only critical aspect addressed by Utescher et al. (2014: 61) regarding the completeness of the fossil record is taphonomy:

> "*Because the Coexistence Approach relies only on the presence or absences of taxa and not on their abundances it is largely independent of sampling size or sampling intensity and can be applied on all plant fossil organs, and where the assemblage is diverse, the CA is* ***robust to taphonomic filtering****.*" [side note: this has never been tested or documented and, hence, is an unfounded claim]

We do not believe that any of this qualifies for "*duly acknowledged*". In fact, the purposes of the 2014 paper seems to be to pick up a number of the critiques by Grimm and Denk (2012) as their own (Grimm et al., 2015, table 1), insinuating that gridded distribution/climate data (in particular, the widely used WorldClim data, Hijmans et al., 2005) is inferior to the "*careful selection of 4-6 stations*" covering only 30 years of climate surveillance (Utescher et al., 2014, fig. 3), and provide a demonstration of power to silent or, at least, channel upcoming open critique: "*With more than 15 years of experience of using the Coexistence Approach on Cenozoic micro- and macropalaeobotanical records, it is clear that the method provides a robust palaeoclimatic proxy.*" (Utescher et al., 2014: 69)

SC4: *§3 Species distribution modelling*

The authors' idea on how SDM could actually help to reconstruct palaeoclimates is not clear. How could modelled (hence based on their own assumptions) n-dimensional niches be used in that specific task? The idea should be detailed. As it stands now, the entire section seems superfluous.

R11: The title of this section is "*Lessons to be learnt from species distribution modelling*". We don't write anything along the line mentioned above. The lessons to be learnt are:

1. If one cannot capture the niche of present day species using rectilinear envelopes, how should this be unproblematic for reconstructing palaeoclimate for fossil assemblages?
2. If the realised niche changes within very short periods of time, how can we assume that the nearest-living-relative principle will result in accurate and precise reconstructions?

However, there are certainly avenues of SDM research that could be used in a mutual climate range setting - i.e. the niche overlap statistics. We have highlighted these in the revised manuscript.

Figures:

***SC5:** All the figures are of good general quality. Figure 2 and/or its caption could however be improved to make it more easily understandable.*

R12: Please could we have more specifics.

Conclusion:

***SC6:** This paper discusses many interesting questions regarding the coexistence approach, but most of those facts are already known by the community.*

R13: If most of these facts are known by the community, it is a wonder that the method is still being used by such a large group of researchers, and uncritically. Please see **R2** and **R9**.

SC7: *The manuscript lacks of 1) a strong discussion and 2) the proposition of alternative methods and/or approaches if we are to, as the authors propose, disregard the CA and all the results it has produced during the last two decades.*

R14: Please see **R2**. The purpose of the manuscript is to stop the usage of a flawed method. Accepting this, will encourage investigation of alternative methods for palaeoclimate reconstructions. Alternative taxon-based methods are referenced, and it is up to the applicants to use and test them. We would like to stress once more, that if a method is fundamentally flawed and in all likelihood producing erroneous or biased reconstructions, it should not be used, *even if there are no alternatives*.

SC8: *Finally, the paper is too aggressive and should be largely toned down.*

R15: We agree that the paper is aggressive, but do not wish to 'tone it down'. We cannot leave any room for misinterpretation or overlooking of suggestions as has occurred previously between the Grimm and Denk (2012), Utescher et al. (2014) papers, and actual applications (see also **R9**).

SC9: *The manuscript may become acceptable for publication in CP after the authors have addressed the major points raised above. We recommend that the corrections should be reviewed.*

R16: We uploaded the edited manuscript as an attachment to this response which the reviewer(s) is free to provide further commentary, the paper will be open for discussion for another month (13 Feb 2016). We would appreciate answers to **R7** and **R12** in order to be able to elaborate.

Boyle, B., Meyer, H. W., Enquist, B., and Salas, S.: Higher taxa as paleoecological and paleoclimatic indicators: A search for the modern analog of the Florissant fossil flora, Geol. Soc. Am. Sp. Paper, 435, 33–51, 2008.

Grimm, G. W., and Denk, T.: Reliability and resolution of the coexistence approach — A revalidation using modern-day data, Rev. Palaeobot. Palynol., 172, 33–47, 2012.

Grimm, G. W., Bouchal, J. M., Denk, T., and Potts, A. J.: Fables and foibles: a critical analysis of the Palaeoflora database and the Coexistence Approach for palaeoclimate reconstruction., bioRxiv, doi, 10.1101/016378, 2015.

Grimm, G. W., and Potts, A.: Fallacies and fantasies: the theoretical underpinnings of the Coexistence Approach for palaeoclimate reconstruction, Climates of the Past Discussions, 11, 5727–5754, 2015.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A.: Very high resolution interpolated climate surfaces for global land areas, Int. J. Climat., 25, 1965-1978, 2005.

Huang, Y.-J., Chen, W.-Y., Jacques, F. M. B., Liu, Y.-S. C., Utescher, T., Su, T., Ferguson, D. K., and Zhou, Z.-K.: Late Pliocene temperatures and their spatial variation at the southeastern border of the Qinghai–Tibet Plateau, J. Asian Earth Sci., doi, 10.1016/j.jseaes.2015.1004.1048, 2015.

Mosbrugger, V., and Utescher, T.: The coexistence approach -- a method for quantitative reconstructions of Tertiary terrestrial palaeoclimate data using plant fossils, Palaeogeogr. Palaeoclimat. Palaeoecol., 134, 61-86, 1997.

Utescher, T., Bruch, A. A., Erdei, B., François, I., Ivanov, D., Jacques, F. M. B., Kern, A. K., Liu, Y.-S. C., Mosbrugger, V., and Spicer, R. A.: The Coexistence Approach—Theoretical background and practical considerations of using plant fossils for climate quantification, Palaeogeogr. Palaeoclimat. Palaeoecol., 410, 58–73, 2014.

Utescher, T., Bondarenko, O. V., and Mosbrugger, V.: The Cenozoic Cooling – continental signals from the Atlantic and Pacific side of Eurasia, Earth and Planetary Science Letters, 415, 121–133, 2015.

Fallacies and Fantasies: the theoretical underpinnings of the Coexistence Approach for palaeoclimate reconstruction

Guido W. Grimm[1] and Alastair J. Potts[2]

[1] University of Vienna, Department of Palaeontology, Vienna, Austria

[2] Nelson Mandela Metropolitan University, Centre of Coastal Palaeoscience, Port Elizabeth, South Africa

Abstract

The Coexistence Approach has been used to infer palaeoclimates for many Eurasian fossil plant assemblages. However, the theory that underpins the method has never been examined in detail. Here we discuss acknowledged and implicit assumptions, and assess the statistical nature and pseudo-logic of the method. We also compare the Coexistence Approach theory with the active field of species distribution modelling. We argue that the assumptions will inevitably be violated to some degree and that the method has no means to identify and quantify these violations. The lack of a statistical framework makes the method highly vulnerable to the vagaries of statistical outliers and exotic elements. In addition, we find numerous logical inconsistencies, such as how climate shifts are quantified (the use of a "center value" of a coexistence interval) and the ability to reconstruct "extinct" climates from modern plant distributions. Given the problems that have surfaced in species distribution modelling, accurate and precise quantitative reconstructions of palaeoclimates (or even climate shifts) using the nearest-living-relative principle and rectilinear niches (the basis of the method) will not be possible. The Coexistence Approach can be summarised as an exercise that shoe-horns a plant fossil assemblages into coexistence and then naively assumes that this must be the climate. Given the theoretical issues, and methodological issues highlighted elsewhere, we suggest that the method be discontinued and that all past reconstructions be disregarded and revisited using less fallacious methods.

Keywords: Cainozoic; Eurasia; mutual climate range; nearest-living-relative principle; physiological uniformitarianism; species distribution modelling; theoretical problems; univariate niche

1 Introduction

One of the most widely used methods to infer the palaeoclimates of Eurasia using fossil plant assemblages is the 'Coexistence Approach' (Utescher et al., 2014). Conceptually, this approach belongs to the family of mutual climate range techniques but also makes use of the 'nearest-living-relative' principle; a nearest-living-relative (NLR) is a modern taxon (species, group of species, genus, or higher) that is considered an analogue for the fossil taxon. Mutual climate range methods use the climatic preferences of modern species (a set of nearest-living-relatives), as defined by their current distribution, to infer the potential climatic niche for a fossil assemblage. In the case of the Coexistence Approach, the climate niche is defined using minimum and maximum climate values of an NLR, obtained from its present-day distribution. Pure mutual climate range techniques are usually restricted to reconstructing palaeoclimates of the recent past (i.e. Quaternary) where species in the fossil assemblages can be directly linked to modern species (e.g. Elias, 1997, 2001; Thompson et al., 2012; Harbert and Nixon, 2015); the processes of extinction and speciation are ignored and niche conservatism is considered to be the norm. However, to apply these palaeoclimate reconstruction techniques to assemblages from older time periods requires the use of the nearest-living-relative principle, which is linked to the concept of physiological uniformitarianism (Tiffney and Manchester, 2001; Tiffney, 2008). The niche-space of an NLR is used to represent that of the fossil taxon. Thus, one assumes that the climate niche of the NLR (the modern species or species set) is identical to that of the associated fossil taxon (an extinct sister or ancestral species) and the mutually shared climate range of the NLRs enables the estimation of the climate conditions in which the fossil assemblage thrived (Fig. 1).

Despite the availability of alternative palaeoclimate reconstruction techniques using NLRs and the mutual climate range approach (e.g. Greenwood et al., 2005), the Coexistence Approach has become the *de facto* method for plant fossil assemblages of Eurasia for time periods spanning the Miocene to Late Cretaceous (Utescher et al., 2014). The cumulative citation count of studies using the Coexistence Approach is in excess of 10,000. On the surface, it reconstructs precise palaeoclimatic conditions (usually reported with a precision of 0.1 °C and 1 mm precipitation per month or year) based on a series of acknowledged and implicit basic assumptions (Table 1; Mosbrugger and Utescher, 1997; Utescher et al., 2014). These assumptions appear straightforward, but have theoretical and practical implications essentially ignored in the application of the Coexistence Approach (Mosbrugger and Utescher, 1997; Utescher et al., 2014; Grimm et al., 2015). Furthermore, the Coexistence Approach avoids any statistical processing (Mosbrugger and Utescher, 1997; Utescher et al., 2014). It relies to some degree on hard-to-grasp pseudo-logic, some of which is advocated as strengths of the method, e.g. the ability to reconstruct extinct climates (Utescher et al., 2014). The applicability of the nearest-living-relative principle for reconstructing past climates in a quantitative manner is never questioned. This is surprising in the light of ongoing discussions in the field of spatial distribution modelling, which shares a number of assumptions with mutual climate range and nearest-living-relatives methods. Below we discuss each of these issues in further detail.

2 Theoretical background of the 'Coexistence Approach'

2.1 Assumptions of the Coexistence Approach

Mosbrugger and Utescher (1997) list four basic assumptions that need to be fulfilled (Table 1). The first assumption has never been used in the application of the Coexistence Approach, and the three others superimpose additional uncertainty on the method and are easily violated, particularly if the aim is high accuracy *and* precision. Notably, none of the assumptions have been tested and verified for taxa commonly used in the Coexistence Approach.

The first assumption is anchored on the ability to define a "systematically close" NLR (Table 1). However, Mosbrugger and Utescher (1997) or Utescher et al. (2014) do not provide a framework on how to quantify "systematically close" and in what respect systematic closeness should be relevant for the identification of the NLR. A focus on "systematic closeness" can lead to conflict with the nearest-living-relative principle. This principle is based on overall morphological similarity and not necessarily linked to phylogenetic relatedness, which is the current basis of systematics. Thus, a fossil may be "systematically close" to a modern species (or group) that has undergone significant shifts in morphology and fundamental niche, and the best modern analogue may be a more distantly related lineage that has been morphologically and ecologically stable (Fig. 2A). In addition, the degree of systematic relatedness of a fossil to an NLR requires the placement of fossils within a phylogenetic framework (i.e. a tree or network) and this has never been explored in any Coexistence Approach study.

There are further issues with Assumption 1 when considering the taxonomic affiliation of an NLR. Given the timespan separating ancient assemblages and modern day taxa, it has been agreed that defining an NLR at the species level is highly problematic (Grimm and Denk, 2012; Utescher et al., 2014). Thus, the Coexistence Approach usually defines a “systematically close” NLR as the genus or family to which the fossil can be assigned, with rare instances of an intrageneric lineage or a modern species (Grimm and Denk, 2012; Utescher et al., 2014; Grimm et al., 2015). For example, the NLR of a fossil oak leaf would be genus *Quercus*, the NLR of a deciduous, convexly lobed oak leaf would be *Quercus* Group Quercus (the white oak clade), and the NLRs of a fagaceous fossil of unknown generic affinity would be all Fagaceae. Hence “systematically close”, as used in the Coexistence Approach and other nearest-living-relative approaches, translates into simply being a member of the same taxonomic rank (e.g. genus or family), and the actual phylogenetic (= systematic) distances between fossils and their NLRs is never established. Under this implementation of assigning NLRs to higher taxonomic ranks (above species) includes the taxonomic problems linked to paraphyly (exclusive common origin; Fig. 2B). Fossils of a paraphyletic group will have different systematic distances to the modern members of the specified taxonomic group of NLRs. However, this is not a problem for the combination of mutual climate range approaches and nearest-living-relative principle as long as the assumption of physiological uniformitarianism is fulfilled (**Assumption 2**). Thus, shared ancestry remains important, but the ‘systematic closeness’ of **Assumption 1** is entirely superfluous for the application of mutual climate range techniques making use of the nearest-living-relative concept.

The second assumption (Table 1) is based upon the concept of physiological uniformitarianism (Tiffney and Manchester, 2001; Tiffney, 2008). Physiological uniformitarianism implies that as long as lineage stays within its environmental niche, it will not accumulate morphological changes. Hence, a modern species with the same, or very similar, morphological traits of a fossil of the same evolutionary lineage should share the same environmental niche. It also implies that members of the lineages that have undergone niche shifts also experienced morphological changes. **Assumption 2** is likely to be violated when morphological changes are evident between the fossil and modern members of an evolutionary lineage and an NLR of a fossil specimen should only be used if there is morphological, not mere taxonomic, similarity (ideally identical) and have a common origin. This would exclude the use of most modern plant genera and all families as NLRs as they are typically composed of morphologically divergent species.

In addition, the use of morphologically diverse taxonomic groups to represent an NLR usually means that the environmental niche of the NLR is large, likely encompassing the niche of the fossil, but is not "climatically similar" to that of the fossil; thus, directly violating **Assumption 2.** Novel procedures and methods are required that take cognisance of the fact that the NLR niche is likely to be far broader than can be expected for that of the fossil. The actual assumption, as used by Coexistence Approach practitioners, is that *the climatic niche of a fossil taxon lies somewhere within the range of niches found within the species comprising the NLR*. This has two major implications for the setup and interpretation of reconstructed palaeoclimates using the Coexistence Approach (and other mutual climate range techniques that use NLRs): 1) a high resolution climate reconstruction should not be possible, especially when only minimum and maximum NLR tolerances are used (Fig. 3A), and 2) mixed floras may not be identified as mutually exclusive species (or communities) can have overlapping climate ranges at higher taxonomic levels (Fig. 3B). Thus, highly precise *and* accurate climate reconstructions can only be obtained using the Coexistence Approach if the critical species within a palaeoassemblage occupied niches close to the minimum **and** maximum tolerances of their corresponding modern genus- or family level NLRs.

The third assumption (Table 1)**,** that the distributions of extant species are in equilibrium with their climate, is a topic rich with discussions in the ecological and species distribution modelling literature (Araújo and Pearson, 2005; Bond et al., 2005; Sexton et al., 2009; Franklin, 2010). Species are often not in equilibrium with their climate for abiotic (e.g. soil, fire) or biotic (e.g. competition) reasons, and thus their realised niches do not span their fundamental niches. Thus species will be plastic in their expression of the realised niche depending on external factors, which would exclude the reconstruction of palaeoclimate with high accuracy. Any change in the abiotic or biotic parameters can affect the distribution of a species (i.e. its realised niche) even if the fundamental niche remains unchanged.

The climatic niche is solely represented by minimum and maximum values in the Coexistence Approach, which are independently compiled for climate parameters in a univariate manner. However, it has been long established that biological climate niches are multi-dimensional (Köppen, 1936; Hutchinson, 1957; Walter, 1973; Walter and Breckle, 1983–1991; Schroeder, 1998). Using minimum-maximum tolerances along univariate axes can only roughly approximate the multidimensional climatic niche, and may be misleading (Klotz, 1999; Thompson et al., 2012). For example, two *mutually exclusive* taxa, for which **Assumption 3** applies, may still have an artificial mutual climate range regarding their minimum and maximum tolerances (Fig. 4A). In this context it is important to note that species distribution modelling started with algorithms that used minimum and maximum values, but quickly moved on to methods that better represented the bioclimatic niche of a species (discussed further below). Thus, the use of range values for climatic parameters does not accurately capture the climatic requirements or tolerances of an NLR (Table 1), which will affect the reconstructed palaeoclimate using the Coexistence Approach.

The fourth and last assumption has no apparent theoretical implications. Technical implications have been discussed in Grimm and Denk (2012), Thompson et al. (2012), Utescher et al. (2014), and Grimm et al. (2015). We do, however, wish to highlight that since local climate can substantially vary over short time scales, minimum and maximum tolerances may be unduly affected by the selected observation period of climate stations.

Not formally addressed by Mosbrugger and Utescher (1997) or Utescher et al. (2014) are two more fundamental assumptions ~~in the application~~ of the Coexistence Approach, which distinguish the method from mutual climate range techniques using modern-day species: 1) palaeoassemblages comprise only taxa that existed as actual communities (i.e. all fossil specimens are autochthonous and from the same point in time); 2) absence of a fossil taxon indicates true absence (i.e. each fossil plant assemblage comprehensively reflects the actual palaeocommunity; Table 1). The Coexistence Approach implicitly assumes that only an autochthonous and strictly coeval palaeoassemblage will result in a single coexistence interval. However, given that two mutually exclusive taxa can share a climate range of minimum and maximum along univariate climate parameters, so too can allochthonous taxa in a fossil assemblage. In addition, the expansion of the climate niche using higher-level NLRs automatically increases the probability of artificial coexistence. Thus, allochthonous assemblages (mixed floras) do not necessarily result in ambiguous intervals (Fig. 4B-D) and may very well be the reason for highly precise palaeoclimate estimates (< 1 °C for temperature parameters, < 100 mm precipitation per year, < 10 mm precipitation per month) observed in many studies using the Coexistence Approach (Denk et al., 2012; Grimm et al., 2015). Thompson et al. (2012) suggest that a benefit of mutual climate range techniques, in comparison to indicator-species approaches, is that the reconstruction is only affected by the presence of taxa, not their absence. However, this does not apply to the Coexistence Approach, where the mere absence of a taxon can directly affect the outcome of the reconstruction (discussed further below). For instance, absence of a taxon may eliminate another NLR as "climatic outlier" rather than producing two "ambiguous" intervals.

We have outlined a range of probable and inevitable issues of the purported basic assumptions of the Coexistence Approach. These will all, to some unknown degree, decrease the precision and accuracy of any approach that attempts to reconstruct palaeoclimates. In this light, the Coexistence Approach is highly unlikely to reconstruct precise or accurate palaeoclimatic conditions. Utescher et al. (2014) state that it is impossible to test the accuracy of Coexistence Approach reconstructions (but see Grimm and Denk, 2012, for mean annual temperature estimates), but follow the original paper in assuming that violation of the basic assumptions will readily surface in the form of "*climatic outliers*".

2.2 The statistical nature of the Coexistence Approach

According to Utescher et al. (2014) the "*Coexistence Approach by Mosbrugger and Utescher (1997) is a nearest living relative method, which relies only on the presence/absence of a plant taxon within a fossil assemblage and the climatic requirements of its modern relatives. It avoids any statistical processing or further assumptions, except those given in Mosbrugger and Utescher (1997)* [i.e. the four basic assumptions, see Table 1]". In the original paper, no means of statistical processing were proposed, hence, the Coexistence Approach defines an interval for a past climate parameter assuming that statistical effects do not exist or are negligible. The Coexistence Approach ignores the majority of the community information because the reconstructed climate interval is always solely defined by the pair of the two most divergent, but putatively coexisting NLRs. Usually one member of the pair is an exotic element; here we define 'exotic' as any NLR whose niche is at odds with the majority of the assemblage (e.g. Fig. 5). The Coexistence Approach lacks a statistical framework to account for potential oddities, errors or violations of assumptions, the likelihood of which increases with assemblage size or depositional age. The approach naively relies on the presumption that any violation will readily surface in the form of so-called "*climatic outliers*" (Mosbrugger and Utescher, 1997; Utescher et al., 2014). This exposes palaeoclimate reconstructions using this approach to the vagaries of statistical outliers and exotic elements (see Grimm and Denk, 2012; Grimm et al., 2015, for real-world data examples).

A 'climatic outlier' is identified as an NLR or small number of NLRs that do not share the climate space for a given parameter with a slightly higher number of other NLRs (Fig. 5). In those cases where there are more than one interval that can be reconstructed using the same maximum possible number of NLRs, then alternative '*ambiguous*' intervals are reported; each of these intervals recognises a different set of climatic outliers. Ambiguous intervals are interpreted by Utescher et al. (2014) as the only evidence for mixed floras rather than a violation of any the assumptions discussed above. Taxa identified as climatic outliers are typically removed from a Coexistence Approach analysis for a given palaeoassemblage and parameter. We wish to highlight that a climatic outlier is simply an NLR that is seemingly at odds with a few other NLRs and must not to be confused with a statistical outlier (Fig. 5).

There are two paramount problems with the current outlier elimination strategy used by the Coexistence Approach. First, *two* taxa *violating* the assumptions behind the Coexistence Approach may eliminate *one* taxon that is not. A typical situation is illustrated in Figure 5, where an NLR occupying a climate range that is in general agreement with the rest of the flora would be identified and eliminated as a climatic outlier because of presence of two deviant taxa that are at odds with the overall NLR community. Second, taxa identified as climatic outliers for one climatic parameter and therefore removed from the assemblage for estimating that parameter are still, in most cases, kept for analysing other parameters for the same assemblage. In some cases, these climatic outliers even define the coexistence interval in another parameter (Grimm et al., 2015). If we follow the logic that climatic outliers represent violations of the basic assumptions of the Coexistence Approach (Utescher et al., 2014), then it is imperative that they are removed from all reconstructions for a given assemblage or in general (Table 2). This has been rarely applied in any study that has identified climatic outliers in the Coexistence Approach, mainly to avoid wide, and thus uninformative, coexistence intervals (Grimm and Denk, 2012; Grimm et al., 2015). It could be argued that any palaeoassemblage represented by mutually exclusive NLRs should be ignored until the reason for the non-coexistence can be identified and corrected for.

Any mutual climate range approach needs a framework to identify statistical outliers as the assumptions will inevitably be violated, and establishing the degree of violation (e.g. degree of niche shifts) is not feasible based on current knowledge. Many palaeoassemblages will comprise mixed floras with elements from different climate niches, and this would need to be explicitly addressed before reconstructing coexistence intervals. As stated above, the Coexistence Approach lacks any framework to identify exotic elements or allochthonous assemblages, unless they are sufficiently divergent to generate climatic outliers. Allochthonous assemblages comprising mutually exclusive species can share a climate interval (Fig. 4B), and this problem of pseudo-coexistence is exacerbated by the use of *higher-level taxa* (genera, families) as NLRs of a fossil *species/morphotypes*. Any slightly conflicting, but exotic, element in an assemblage will have a disproportionally high influence on the palaeoclimate estimates (Fig. 5). It is clear that not only 'climatic outliers' and 'ambiguous intervals' should be indicative of mixed floras, errors in the data, or violations in the assumptions, but also *any* narrow coexistence interval (see Grimm and Denk, 2012; Grimm et al., 2015, for real-world examples).

Mutual climate range techniques that apply simple statistics to filter exotic taxa, such as the Bioclimatic Approach (Greenwood et al., 2005), will be less susceptible to the presence and absence of a few exotic taxa, but will also usually fail to recognise mixed floras. The problem of mixed floras can only be overcome, to some degree, by using alternative mutual climate range techniques that make use of the full spectrum of distributional information, and thus include the climatic preference of all constituent elements of a palaeoassemblage (e.g. using the niche curves in Fig. 5). This includes methods such as the weighted mutual climate range approach (Thompson et al., 2012), the probability density function method (Chevalier et al., 2014), and the coexistence likelihood estimation method (Harbert and Nixon, 2015). However, these methods will probably begin to break down when the nearest-living-relative principle is needed to link fossils with extant lineages (Thompson et al., 2012; Harbert and Nixon, 2015), and may explain why their application has been limited to Quaternary assemblages.

2.3 Pseudo-logic of the Coexistence Approach

We wish to highlight four additional points regarding the use of the Coexistence Approach that lack any (bio)logical basis, specifically: 1) the use of the "center value" to identify and quantify climatic shifts, 2) that the reconstructed climate is based on only two nominally coexisting elements, 3) that the reconstructed climate is highly dependent on the presence or absence of a single or few taxa (the "Heisenberg effect"), and 4) the reconstruction of "extinct" climates. We elaborate on each of these points below.

The conclusions of most Coexistence Approach studies rely on shifts observed in the so-called "center value". This value is simply the arithmetic mean of the upper and lower boundary of the coexistence interval. Practical tests have shown that there is little correlation between the actual climate and the "center value" (Klotz, 1999; Grimm and Denk, 2012). The use of this value highlights a fundamental misunderstanding of the niche concept. If we imagine the coexistence interval to be correct, then ***all*** values within the interval should be equally probable as no other information is incorporated regarding the probabilities of occurrence of the assemblage. Selecting the "center value" as an indicator of a shift in climate makes no statistical or biological sense. For example, Figure 6A shows two plant assemblages that differ only by the climatic preference of a single NLR. The replacement of one NLR by another with a preference towards *lower* values gives rise to a reconstructed climate shift towards *higher* values using the "center value".

Many Coexistence Approach reconstructions rely on the presence of NLRs that nominally coexist, even if these elements have climate tolerances that are at odds with the rest of the assemblage (Fig. 5; cf. Grimm and Denk, 2012; Grimm et al., 2015). In extreme cases the same coexistence interval can be reconstructed based on plant assemblages with contrary climate tolerances (Fig. 6B). In Figure 6B, the elements of two plant assemblages have contrary climate tolerances and it is the two exotic taxa in each assemblage that ensure that the reconstructed coexistence intervals are the same. Thus, the precision of the reconstructed palaeoclimates is often entirely dependent on the presence or absence of specific, usually exotic NLRs. Across Coexistence Approach studies, a handful of NLRs that occur towards the tolerance margins over the *entirety* of all palaeofloras usually determine the coexistence intervals; it is these few NLRs that give rise to the praised precision of the technique (Grimm and Denk, 2012; Grimm et al., 2015).

The presence or absence of individual NLRs are generally at the base of reconstruction uncertainty in coexistence interval - we term this the "Heisenberg" effect. Figure 6C shows two very similar assemblages where the presence or absence of the two highlighted taxa changes the coexistence interval reconstructed by the Coexistence Approach in a degree that would be interpreted as a trend towards higher values. The Heisenberg effect renders palaeoclimate estimates obtained with the Coexistence Approach protocol highly susceptible to taxon-bias effects. The reconstructed climate is exceedingly dependent on what fraction of the actual vegetation has been captured by the fossil assemblages (note that in Fig. 6C all

NLRs have a mutually shared climate range). Thus, even if all assumptions needed for a mutual climate range approach that also uses the nearest-living-relative principle are fulfilled, the Heisenberg effect will lead to unstable, even random, climate reconstructions.

Utescher et al. (2014) explicitly state that, as each parameter is independently reconstructed, the Coexistence Approach has the potential to reconstruct a climate that does not exist today: an “extinct climate”. It is hard to grasp how this can be logically accommodated with the basic assumptions of the Coexistence Approach and the actuo-palaeontological nearest-living-relative principle in general (Fig. 7). An extinct climate for a palaeoassemblage would indicate that the present-day niches of the NLRs are **not** representative of the fossils, and therefore would indicate direct violations of Assumptions 2 and 3 discussed above (Table 1). In addition, it is not possible to construct an extinct climate using species that are restricted to present-day climates if the principle of physiological uniformitarianism applies. Reasons why extinct climates are reconstructed using the Coexistence Approach include violations of basic assumptions, pseudo-coexistence, the inconsistent identification of climatic outliers within an assemblage across climate variables, and the single-dimension effect where climate parameters are analysed in isolation and are assumed to be unlinked. The reconstruction of an extinct climate should be seen as a direct indication of error, and not lauded as a benefit of the method.

Ignoring these logical inconsistencies in the conception and application of the method, the Coexistence Approach still cannot be expected to reproduce a robust quantitative reconstruction of the palaeoclimate, as 1) assumptions are likely to be violated but cannot be detected, 2) one cannot avoid using higher-level taxa to represent fossil species or morphotypes, and 3) the fossil record will always be incomplete to different degrees, and this will affect the calculated coexistence interval.

3 ~~Applicability of the nearest-living-relative principle: l~~Lessons to be learnt from species distribution modelling

Species distribution modelling (SDM) is an exceptionally active field which aims to empirically model the species-environment relationships and thereby quantify the realised niche of a given taxon (Franklin, 2010; Peterson et al., 2011) or, in some cases, communities (e.g. Potts et al., 2013). The dawn of the field was the BIOCLIM software package (Nix, 1986), which is comparable to the Coexistence Approach as it used the range (or percentile range) of climatic variables in a rectilinear fashion. Booth et al. (2014) describe the roots of the field and highlight that one of the most active areas of SDM development has been of methods that trim the rectilinear climate envelopes of BIOCLIM. This was driven by the early realisation that the relationships between climate variables ~~was~~ were poorly captured by the rectilinear approach; for example, a rectilinear niche may suggest that a species could survive in a situation where it is both hot and dry, but the actual climate niche is that it only occurs where it is hot and wet. More advanced methods have refined the *n*-dimensional hyperniche (Hutchinson, 1957) where response curves are used to capture the suitability of different conditions for species occurrence. ~~This revolution in the multidimensional quantification of the niche has completely bypassed the Coexistence Approach and many other nearest-living-relative methods.~~ BIOCLIM performed poorly in comparison to more recent methods in a comparison of more recent SDM methods (Elith et al., 2006) indicating that the simplistic use of range values for climatic variables, as used by the Coexistence Approach, is a poor representation of the realised niche of species or NLR.

The revolution in the multidimensional quantification of the niche has completely bypassed the Coexistence Approach. Measuring ecological niche overlap between species in multivariate space is an active area of investigation (Rödder and Engler, 2011; Broennimann et al., 2012), which can be used to determine the shared niche within a set of species. However, measuring the niche in such a manner also requires that all the variables selected are, in fact, significant in limiting the niche. Establishing the contribution and importance of different environmental variables (i.e. variable selection) in setting the bounds of a taxon’s niche is a theoretical issue (Araújo and Guisan, 2006) where advances are also being made (Austin and Van Niel, 2011). In comparison, the Coexistence Approach blindly uses a wide range of environmental parameters in a univariate manner assuming that they are all important in determining a taxon’s niche.

Furthermore, ~~T~~the assumption of niche conservatism (linked to the principle of physiological uniformitarianism) has generated considerable debate in the SDM literature as it has been used as justification for projecting models into altered climate states (past or future) and to predict the establishment and spread of invasive species (reviewed in Pearman et al., 2008a). These discussions have centred firstly on whether the current distribution for a given species, i.e. the realised niche, adequately represents the fundamental niche; and secondly, how quickly the fundamental niche might be able to shift? Such concerns are absent in the theoretical underpinnings of Coexistence Approach (Mosbrugger and Utescher, 1997; Utescher et al., 2014). Unfortunately, niche shifts have been documented for a wide range of plant species ~~both~~ through space (Broennimann et al., 2007; Pearman et al., 2008a) and even over relatively short time scales (Pearman et al., 2008b; Veloz et al., 2011). Therefore the assumption of physiological uniformitarianism has limited applicability to reconstruct precise *and* accurate palaeoclimates, especially with increasing age of an assemblage.

4 Conclusions

Using best-possible climate data for modern North American woody plants, Thompson et al. (2012) were unable to reconstruct the climatic shifts from the Last Glacial Maximum to the present-day using an unweighted mutual climatic range method (which represents the niche using range values and is equivalent with the Coexistence Approach save for the use of NLRs and recognition of climatic outliers). This is in stark contrast to the beliefs of Coexistence Approach practitioners that the method can reconstruct climate shifts at high-precision, despite the additional error and uncertainty associated with the nearest-living-relative principle. The purported high precision in Coexistence Approach studies is dependent on phenomena such as pseudo-coexistence and the lack of a sound statistical framework.

We argue that the Coexistence Approach, as conceived by Mosbrugger and Utescher (1997), violates the basic concepts behind mutual climate range techniques and the nearest-living-relative principle. It imposes a number of assumptions that will inevitably be violated and has no ability to detect violations and generally lacks any safeguards against the reconstruction of artificial coexistence intervals and thus erroneous palaeoclimate estimates.

Given the theoretical problems outlined here, and the practical problems highlighted by Grimm et al. (2015) – for example, that any random real-world flora will eventually produce a “statistically significant” (according Mosbrugger and Utescher, 1997) coexistence interval with a high number (>20) of NLRs – we suggest that all palaeoclimate reconstruction studies using the Coexistence Approach be disregarded and that the palaeoassemblages be revisited with other methods and careful, well-documented, and well-investigated NLR-associations.

5 Where to go from here?

There are already a range of potential methods available for palaeoclimate reconstruction using plant fossils as proxies in a univariate manner that have been rarely used or recently proposed (e.g. Greenwood et al., 2005; Boyle et al., 2008; Thompson et al., 2012; Chevalier et al., 2014; Harbert and Nixon, 2015), and there are avenues ripe for exploration (e.g. Broennimann et al., 2012; Denk et al., 2013). However, all of these methods require (further) testing, and then careful, well-documented usage when reconstructing palaeoclimates. The development of the physiognomic approach (CLAMP) within the last two decades may serve as an example regarding validation, advancement, and, most importantly, documentation and transparency. The various publications demonstrate a constant thrive to reach higher precision and counter known problems (e.g. Kovach and Spicer, 1995; Herman and Spicer, 1997; Spicer et al., 2009; Yang et al., 2015; Li et al., 2016). All primary data is made freely accessible and means are implemented allowing for quick application (CLAMP online; Yang et al., 2011). CLAMP online does not only provide data, guidelines and templates for application, but also pinpoints shortcomings and ideas how to deal with them. No method is or will be ‘perfect’. Nevertheless, it is crucial to define the principal accuracy and precision of any quantitative method. If this is not possible, as in the case of the Coexistence Approach (Utescher et al., 2014: 61), it must not be used. Therefore, we suggest that any current or future taxon-based method be:

1. Tested against the modern flora (e.g. Boyle et al., 2008; Thompson et al., 2012; Chevalier et al., 2014; Harbert and Nixon, 2015).
2. Tested with randomised and unlikely communities of modern flora. A robust (taxon-based) method that is to be applied to micro-, meso- and macrofossil assemblages must detect possible allochthonous elements/mixed floras.
3. First applied to the ‘better-understood’ palaeoclimates of the most recent past (e.g. present to the Last Glacial Maximum) and compared with available relevant proxies (e.g. Thompson et al., 2012).
4. Explored using both species-level and taxonomic levels of potential or probable nearest-living relatives (e.g. Boyle et al., 2008).
5. Examined using a jackknifing or similar procedure to ensure that results remain accurate and establish the actual precision that can be expected with fossil floras. Fossil floras will always only provide a fraction of the actual flora, and may include

incorrectly determined taxa. The accuracy of a result must not change due to the presence or absence of specific taxa in the assemblage, although precision can, and is likely to, decline.

6. Finally, tested in a stepwise fashion further and further into the past using available, well-studied and dated, more or less continuous records, such as the recently revised Icelandic record covering the last 15 million years ranging from subtropical lowland to ice age conditions (Denk et al., 2011; Denk et al., 2013)

After such a series of tests, the method can be considered an alternative means to reconstruct past climates for further exploration. However, the ultimate limitations of mutual climate range techniques or other nearest-living-relative methods for palaeoclimate reconstruction do not lie in the methodological framework to estimate, for example, the coexistence space, but rather the applicability of the nearest-living-relative principle. When it comes to application into the more distant past, the basic assumption of any method must be that the nearest-living-relative principle will be violated to an unknown degree. The degree of violation will likely increase with time, and may not necessarily surface during application or testing phase. Bivariate or multivariate approaches, which can tackle the problem of pseudo-coexistence (e.g. Fig. 4), will be more sensible in this respect. The capability to accurately and precisely predict palaeoclimate will not only deteriorate with increasing age, but also with compositional change of the fossil plant assemblages relative to the modern-day situation. Precise, highly sophisticated methods (e.g. Punyasena, 2008; Harbert and Nixon, 2015) or methods using few, overly precise, values to characterise the niche space of the NLR (e.g. Greenwood et al., 2005) run a higher risk of being affected by violations of the nearest-living-relative principle than methods that use semi-quantitative approximations of the niche (e.g. Thompson et al., 2012; Denk et al., 2013).

Taking into account all theoretical and practical issues involved, we suspect that quantitative palaeoclimate estimates at a high precision *and* accuracy is an impossible goal when the nearest-living-relative principle has to be applied. Therefore, our opinion is that method development should not focus on high (or higher) precision, as the basis of this precision is undermined as the temporal difference between fossil and NLR increases, but rather on establishing climate change trends in a robust and reproducible manner. Semi-quantitative approaches can detect such changes and may prove to be more robust (e.g. the Köppen signature approach proposed by Denk et al., 2013). Furthermore, the application of any nearest-living-relative method on palaeofloras will always depend on thoughtful filtering a fossil assemblage for elements that have been shown to have a high likelihood of niche conservatism. Fossil-NLR associations must be carefully selected to ensure that the principle of physiological uniformitarianism applies, in contrast to the current practise of seemingly data-naive bulk analyses.

Acknowledgements

This study was funded by the Austrian Science Fund (FWF) with a grant to GWG, project number M1751-B16. AJP received support from the National Research Foundation (RCA13091944022).

References

Araújo, M. B., and Pearson, R. G.: Equilibrium of species' distributions with climate, Ecography, 28, 693–695, 2005.

Araújo, M. B., and Guisan, A.: Five (or so) challenges for species distribution modelling, J. Biogeogr., 33, 1677–1688, 2006.

Austin, M. P., and Van Niel, K. P.: Improving species distribution models for climate change studies: variable selection and scale, J. Biogeogr., 38, 1–8, 2011.

Bond, W. J., Woodward, F. I., and Midgley, G. F.: The global distribution of ecosystems in a world without fire, New Phytol., 165, 525–538, 2005.

Booth, T. H., Nix, H. A., Busby, J. R., and Hutchinson, M. F.: BIOCLIM: the first species distribution modelling package, its early applications and relevance to most current MaxEnt studies, Divers. Distr., 20, 1–9, 2014.

Boyle, B., Meyer, H. W., Enquist, B., and Salas, S.: Higher taxa as paleoecological and paleoclimatic indicators: A search for the modern analog of the Florissant fossil flora, Geol. Soc. Am. Sp. Paper, 435, 33–51, 2008.

Broennimann, O., Treier, U. A., Müller-Schärer, H., Thuiller, W., Peterson, A. T., and Guisan, A.: Evidence of climatic niche shift during biological invasion, Ecology Letters, 10, 701–709, 2007.

Broennimann, O., Fitzpatrick, M. C., Pearman, P. B., Petitpierre, B., Pellissier, L., Yoccoz, N. G., Thuiller, W., Fortin, M.-J., Randin, C., Zimmermann, N. E., Graham, C. H., and Guisan, A.: Measuring ecological niche overlap from occurrence and spatial environmental data, Global Ecol. Biogeogr., 21, 481–497, 2012.

Chevalier, M., Cheddadi, R., and Chase, B. M.: CREST (Climate REconstruction SofTware): a probability density function (PDF)-based quantitative climate reconstruction method, Clim. Past, 10, 2081–2098, 2014.

Denk, T., Grímsson, F., Zetter, R., and Símonarson, L. A.: Late Cainozoic Floras of Iceland: 15 Million Years of Vegetation and Climate History in the Northern North Atlantic., Topics in Geobiology, Springer, Heidelberg, New York, 854 pp., 2011.

Denk, T., Grimm, G. W., and Röseler, A.-K.: When field botany meets history: taxonomy of *Platanus mexicana* in Mexico, Willdenowia, 42, 99–115, 2012.

Denk, T., Grimm, G. W., Grímsson, F., and Zetter, R.: Evidence from "Köppen signatures" of fossil plant assemblages for effective heat transport of Gulf Stream to subarctic North Atlantic during Miocene cooling, Biogeosci., 10, 7927–7942, 2013.

Elias, S. A.: Review of Palaeoentomology of the Kap København Formation, a Plio-Pleistocene Sequence in Peary Land, North Greenland by Jens Böcher, Arctic and Alpine Research, 29, 140, 1997.

Elias, S. A.: Mutual climatic range reconstructions of seasonal temperatures based on Late Pleistocene fossil beetle assemblages in Eastern Beringia, Quaternary Sci. Rev., 20, 77–91, 2001.

Elith, J., Graham, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R. J., Huettmann, F., Leathwick, J. R., Lehmann, A., Li, J., Lohmann, L. G., Loiselle, B. A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J. M., Peterson, A. T., Phillips, S. J., Richardson, K. S., Scachetti-Pereira, R., Schapire, R. E., Soberón, J., Williams, S., Wisz, M. S., and Zimmermann, N. E.: Novel methods improve prediction of species distributions from occurrence data, Ecography, 29, 129–151, 2006.

Franklin, J.: Mapping Species Distributions: Spatial Inference and Prediction, Cambridge University Press, Cambridge, UK, 2010.

Greenwood, D. R., Archibald, S. B., Mathewes, R. W., and Moss, P. T.: Fossil biotas from the Okanagan Highlands, southern British Columbia and northeastern Washington State: climates and ecosystems across an Eocene landscape, Can. J. Earth Sci., 42, 167–185, 2005.

Grimm, G. W., and Denk, T.: Reliability and resolution of the coexistence approach — A revalidation using modern-day data, Rev. Palaeobot. Palynol., 172, 33–47, 2012.

Grimm, G. W., Bouchal, J. M., Denk, T., and Potts, A. J.: Fables and foibles: a critical analysis of the Palaeoflora database and the Coexistence Approach for palaeoclimate reconstruction., bioRxiv, doi, 10.1101/016378, 2015.

Harbert, R. S., and Nixon, K. C.: Climate reconstruction analysis using coexistence likelihood estimation (CRACLE): A method for the estimation of climate using vegetation, Am. J. Bot., 102, 1277-1289, 2015.

Herman, A. B., and Spicer, R. A.: New quantitative palaeoclimate data for the Late Cretaceous Arctic: evidence for a warm polar ocean, Palaeogeogr. Palaeoclimat. Palaeoecol., 128, 227–251, 1997.

Hutchinson, G. E.: Concluding Remarks, Cold Spring Harbor Symp. Quant. Biol., 22, 415–427, 1957.

Klotz, S.: Neue Methoden der Klimarekonstruktion - angewendet auf quartäre Pollensequenzen der französischen Alpen., Tübinger Mikropaläontologische Mitteilungen, Institut & Museum für Geologie & Paläontologie [now: Institute for Geosciences], Eberhard Karls University, Tübingen, 1999.

Köppen, W.: Das geographische System der Klimate., in: Handbuch der Klimatologie, Band 1, Teil C., edited by: Köppen, W., and Geiger, R., Gebr. Borntraeger, Berlin, 1–44, 1936.

Kovach, W., and Spicer, R. A.: Canonical correspondence analysis of leaf physiognomy: a contribution to the development of a new palaeoclimatological tool, Palaeoclimates: Data and Modelling, 2, 125–138, 1995.

Li, S. F., Jacques, F. M. B., Spicer, R. A., Su, T., Spicer, T. E. V., Yang, J., and Zhou, Z.: Artificial neural networks reveal a high-resolution climatic signal in leaf physiognomy, Palaeogeogr. Palaeoclimat. Palaeoecol., 442, 1–11, 2016.

Mosbrugger, V., and Utescher, T.: The coexistence approach -- a method for quantitative reconstructions of Tertiary terrestrial palaeoclimate data using plant fossils, Palaeogeogr. Palaeoclimat. Palaeoecol., 134, 61-86, 1997.

Nix, H. A.: A biogeographic analysis of Australian elapid snakes, in: Atlas of Elapid Snakes of Australia, edited by: Longmore, R., Australian Government Publishing Service, Canberra, 4–15, 1986.

Pearman, P. B., Guisan, A., Broennimann, O., and Randin, C. F.: Niche dynamics in space and time, Trends in Ecology & Evolution, 23, 149-158, 2008a.

Pearman, P. B., Randin, C. F., Broennimann, O., Vittoz, P., van der Knapp, W. O., Engler, R., Le Lay, G., Zimmermann, N. E., and Guisan, A.: Prediction of plant species distributions across six millennia, Ecology Letters, 11, 357–369, 2008b.

Peterson, A. T., Soberón, J., Pearson, R. G., Anderson, R. P., Martínez-Meyer, E., Nakamura, M., and Araújo, M. B.: Ecological Niches and Geographic Distributions, Princeton University Press, Princeton, 2011.

Potts, A. J., Hedderson, T. A., Franklin, J., and Cowling, R. M.: The Last Glacial Maximum distribution of South African subtropical thicket inferred from community distribution modelling, J. Biogeogr., 40, 310–322, 2013.

Punyasena, S. W.: Estimating Neotropical palaeotemperature and palaeoprecipitation using plant family climatic optima, Palaeogeogr. Palaeoclimat. Palaeoecol., 265, 226-237, 2008.

Rödder, D., and Engler, J. O.: Quantitative metrics of overlaps in Grinnellian niches: advances and possible drawbacks, Global Ecol. Biogeogr., 20, 915–927, 2011.

Schroeder, G.-F.: Lehrbuch der Pflanzengeographie., Quelle & Meyer, Wiesbaden, 459 pp., 1998.

Sexton, J. P., McIntyre, P. J., Angert, A. L., and Rice, K. J.: Evolution and ecology of species range limits, Ann. Rev. Ecol. Evol. Syst., 40, 415–436, 2009.

Spicer, R. A., Valdes, P. J., Spicer, T. E. V., Craggs, H. J., Srivastava, G., Mehrotra, R. C., and Yang, J.: New development is CLAMP: calibration using global gridded meteorological data, Palaeogeography, Palaeoclimatolgy, Palaeoecology, 283, 91–98, 2009.

Thompson, R. S., Anderson, K. H., Pelltier, R. T., Strickland, L. E., Bartlein, P. J., and Shafer, S. L.: Quantitative estimation of climatic parameters from vegetation data in North America by the mutual climatic range technique, Quaternary Sci. Rev., 51, 18–39, 2012.

Tiffney, B. H., and Manchester, S. R.: The use of geological and paleontological evidence in evaluating plant phylogeographic hypotheses in the Northern Hemisphere Tertiary., Int. J. Plant Sci., 162, S3-S17, 2001.

Tiffney, B. H.: Phylogeography, fossils, and northern hemisphere biogeography: The role of physiological uniformitarianism., Ann. Missouri Bot. Gard., 95, 135-143, 2008.

Utescher, T., Bruch, A. A., Erdei, B., François, I., Ivanov, D., Jacques, F. M. B., Kern, A. K., Liu, Y.-S. C., Mosbrugger, V., and Spicer, R. A.: The Coexistence Approach—Theoretical background and practical considerations of using plant fossils for climate quantification, Palaeogeogr. Palaeoclimat. Palaeoecol., 410, 58–73, 2014.

Veloz, S. D., Williams, J. W., Blois, J. L., He, F., Otto-Bliesner, B., and Liu, Z.: No-analog climates and shifting realized niches during the late Quaternary: implications for 21st-century predictions by species distribution models, Global Change Biol., 18, 1698–1713, 2011.

Walter, H.: Vegetation of the Earth in relation to the climate and the eco-physiological conditions., Springer Verlag, New York, Heidelberg, Berlin, 237 pp., 1973.

Walter, H., and Breckle, S.-W.: Ökologie der Erde, Eugen Ulmer Verlag, Stuttgart, 1983–1991.

Yang, J., Spicer, R. A., Spicer, T. E. V., and Li, C.-S.: ‘CLAMP Online’: a new web-based palaeoclimate tool and its application to the terrestrial Paleogene and Neogene of North America, Palaeobiodiversity and Palaeoenvironments, 91, 163–183, 2011.

Yang, J., Spicer, R. A., Spicer, T. E. V., Arens, N. C., Jacques, F. M. B., Su, T., Kennedy, E. M., Herman, A. B., Steart, D. C., Srivastava, G., Mehrotra, R. C., Valdes, P. J., Mehrotra, N. C., Zhou, Z. K., and Lai, J. S.: Leaf Form-Climate Relationships on the Global Stage: An Ensemble of Characters, Global Ecol. Biogeogr., 10, 1113–1125, 2015.

Figure legends

Figure 1. The concept of the mutual climate range as used in the Coexistence Approach.

Figure 2. Difference between ‘systematically close’ and nearest-living-relative (NLR, i.e. best modern analogues). Shown is a species phylogeny of a diversified ingroup; the outgroup in this example is a sister species of the ingroup. **A.** Standard definition of nearest-living relative (best modern analogue) vs. definition if Assumption 1 of the Coexistence Approach should be fulfilled. **B.** Same tree as in A, only that each species is categorised as a member of a distinct morphotaxon that can be distinguished in the fossil record. Note that all morphotaxa are mutually exclusive regarding their climatic niche, but there is no strict correlation between systematic closeness (phylogeny) and the climatic niche of the fossils and their nearest-living relatives (modern species of the same morphotypes as the fossils).

Figure 3. Issues related to the use higher-level taxonomic classification (e.g. genus or family) as nearest-living relatives (NLRs) of fossil species. In this example, two fossil species occupy a climate range within the modern climate range of their selected genus-level NLRs, fulfilling the principle of physiological uniformitarianism. **A.** The fossil species have a narrow shared climate range and coexisted in the past. The use of higher-level taxonomic ranks as NLRs will lead, in most cases, to a much broader and less precise reconstructed coexistence interval. **B.** The fossil species are mutually exclusive,

but the expansion of the niche space – due to the use of genera as NLRs – results in a coexistence interval (i.e. pseudo-coexistence).

Figure 4. Pseudo-coexistence as a result of the representation of the climate niche using minimum and maximum tolerances. A. Bivariate climate niches of two mutually exclusive species. These species have no overlapping climate space, but still reconstruct narrow coexistence intervals (orange bars) along univariate axes. **B.** Bivariate climate niches of NLRs of two floras growing under substantially different climates (indicated "x"s). Note that only the niches of three of the Community 1 species overlap with one or two of the Community 2 species. **C, D.** Univariate mutual climate ranges (MCR) of both communities; the overlap of the two MCR result in highly precise coexistence intervals for the artificially mixed communities including all elements from Community 1 and Community 2.

Figure 5. "Climatic outliers" and the bias of the Coexistence Approach towards exotic nearest-living relatives (NLRs). Shown are the niche response curves for 20 potential NLRs, of which 18 (grey and green) show a general overlap in their climatic preference. The two red NLRs are exotic elements with strongly differing climatic preferences. Bars indicate the minimum and maximum tolerances of each NLR, the dots highlight each NLR's optimal climate value. Because the green NLR has no shared climate range with the two exotic NLRs (red), it would be excluded as a "climatic outlier" following the Coexistence Approach protocol. The resultant coexistence interval (orange bar) is highly precise but reflects neither the climatic preference of the non-exotic (grey and green) nor exotic group of NLRs (red).

Figure 6. Logical inconsistencies in the application and theory of the Coexistence Approach. Shown are coexistences intervals (orange bars) based on slightly (**A, C**) or extremely (**B**) different sets of nearest-living relatives (NLRs). **A.** Use of the "center value" to determine climate shifts. A single NLR (black) is replaced by a NLR tolerant to *lower* values (red), which would be eliminated as a "climatic outlier" by the two green NLRs; thus leading to a *higher* "center value"*:* **B.** All NLRs have contrasting climate tolerances, the exotic taxa in both floras (red) ensure that the reconstructed coexistence interval is the same. **C.** Two floras that only differ by the absence (white bars) or presence (black bars) of each a single taxon. The resulting coexistence intervals would be interpreted as a shift towards higher values. The green box shows the coexistence interval of a flora in which both taxa are represented.

Figure 7. Impossibility of reconstructing extinct climates with the nearest-living-relative (NLR) principle. Shown are the (realised) climate niches of five modern species, which, inevitably have to lie within the frame of the modern climate space. Any coexistence space (yellow square, showing the coexistence space of species 2, 3, and 4 using their minimum and maximum tolerances) must reflect a climate situation also found today. Any extinct climate (grey square) could only be defined by the coexistence of species with *different* climate niches than found in modern species, species with no living NLR or species belonging to lineages that underwent niche shift.

Table 1. The assumptions of the Coexistence Approach (quotations from Utescher *et al.*, 2014)

Description	Issues
Assumption 1: "*For fossil taxa systematically closely related nearest living relatives (NLRs) can be identified.*"	a) Lack of a theoretical framework to define a systematically close relative b) Concept of physiological uniformitarianism assumes a common origin, but does not need quantification of phylogenetic closeness
Assumption 2: *"The climatic requirements of a fossil taxon are similar to those of its nearest living relative."*	a) Physiological uniformitarianism cannot be generally assumed b) Different taxonomic ranks of fossils and their nearest-living relatives
Assumption 3: *"The climatic requirements or tolerances of a nearest living relative* [i.e. minimum and maximum tolerances regarding single parameters that are considered per se to be independent from each other] *can be derived from its* [current] *area of distribution"*	a) Distribution is not necessarily a function of climate, but also other biotic and abiotic parameters: the realised niche < fundamental niche b) Minimum and maximum tolerances are poor estimates for the climatic niche of a taxon c) Climate parameters are not independent from each other d) There are no working frameworks to test if a potential nearest-living-relative fulfils **Assumption 3**
Assumption 4: "*The modern climatic data used are reliable and of good quality*"	More or less violated in all studies that applied the Coexistence Approach (see Grimm & Denk, 2012)
Assumption 5: Palaeoassemblages represent actual communities	a) Fossils may be allochthonous, in particular microfossils (pollen). b) Fossils may not be strictly coeval (macrofossil lagerstätten usually cover substantial time periods)
Assumption 6: Absence of a fossil in a palaeoassemblage is evidence of true absence	The fossil record is incomplete

Table 2: The consequences of identifying a 'climatic outlier' in a palaeoassemblage supposing that these represent violations of one or more of the four basic assumptions.

Violation of basic assumption…	**Consequence**
… 1: The nearest-living-relative (NLR) is not a close relative of a fossil taxon	There is no consequence as long as the NLR shares the same lineage and is a good physiological modern analogue.
… 2: The climatic requirements of the fossil taxon are different from the NLR	If different for one climate parameter, the NLR may be equally non-representative for other climate parameters of the fossil taxon. Any coexistence interval including this fossil taxon may be misinformed.
… 3: The NLR's distribution is not representative for its climatic requirements (relict distribution)	Coexistence intervals delimited by the NLR are likely to be misinformed in any study using the NLR.
… 4: The modern climate data to estimate NLR minimum and maximum tolerances are unreliable.	If this is the case, then no coexistence interval is reliable and palaeoclimate reconstruction using modern analogues is impossible.